

Modern Physics Letters A
 Vol. 27, No. 27 (2012) 1230028 (20 pages)
 © CERN
 DOI: 10.1142/S0217732312300285

Brief review of the searches for the rare decays
 $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$

Johannes Albrecht
CERN
1211 Geneva, Switzerland
Johannes.Albrecht@cern.ch

Received 18 July 2012
 Accepted 18 July 2012
 Published 13 August 2012

The current experimental status of the searches for the very rare decays $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ is discussed. These channels are highly sensitive to various extensions of the Standard Model, specially in the scalar and pseudoscalar sector. The recent, most sensitive measurements from the CDF, ATLAS, CMS and LHCb collaborations are discussed and the combined upper exclusion limit on the branching fractions determined by the LHC experiments is shown to be 4.2×10^{-9} for $B_s^0 \rightarrow \mu^+ \mu^-$ and 0.8×10^{-9} for $B^0 \rightarrow \mu^+ \mu^-$. The implications of these tight bounds on a selected set of New Physics models is sketched.

Keywords: Flavour Physics, Rare Decays, Leptonic decays, b-hadron, FCNC, LHC

PACS Nos.: 13.20.He 13.30.Ce 12.15.Mm 12.60.Jv

1. Introduction

Processes, which are highly suppressed in the Standard Model (SM), such as decays mediated by flavour changing neutral currents (FCNC) allow stringent tests of our current understanding of particle physics. These transitions are forbidden at tree level in the SM, as all electrically neutral particles (γ , Z^0 , H^0 and gluons) have only diagonal couplings in the flavor space. FCNC processes are therefore only allowed through loop contributions and probe the underlying fundamental theory at the quantum level, where they are sensitive to masses much higher than that of the b -quark. Historically, many observations have first been indicated by FCNC processes, examples include the existence of the charm quark or the high top quark mass.

Precise measurements of the branching fractions of the two FCNC decays $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ belong to the most promising modes for a possible discovery of a theory beyond the SM. These decays are strongly suppressed by loop and helicity factors, making the SM branching fraction small: 3.1×10^{-9} for B_s^0 decays and 1×10^{-10} for B^0 decays, both known with a precision better than 10%.

Enhancements of the decay rates of these decays are predicted in a variety of different New Physics models. It has been emphasized many times that this decay is very sensitive to the presence of supersymmetric particles^{1,2,3,4,5,6,7,8,9}. For example, in the minimal supersymmetric Standard Model (MSSM), the enhancement is proportional^{9,10,11} to $\tan^6 \beta$, where $\tan \beta$ is the ratio of the vacuum expectation values of the two Higgs fields. For large values of $\tan \beta$, this search belongs to the most sensitive probes for physics beyond the SM which can be performed at collider experiments. Other models such as non minimal flavor violating or Littlest Higgs models as well as those with extra dimensions like Randall Sundrum models predict large effects independent of the value of $\tan \beta$ ^{12,13,14,15,16,17,18,19}. In the absence of an observation, limits on $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$ are complementary to those provided by high p_T experiments. The interplay between both allows to optimally constrain the SUSY parameter space.

Measuring the decay rates of these decays has been a major goal of particle physics experiments in the past decade. The limit on the decay rates was gradually improved by the CDF and D0 experiments at the Tevatron and the CMS, ATLAS and LHCb experiments at the LHC.

In this review, the prediction of the branching fraction of $B_{s,d}^0 \rightarrow \mu^+\mu^-$ is discussed in Sec. 2. The most sensitive measurements of the branching fractions of $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ are discussed in Sec. 3, including a combination of the measurements performed by the three LHC experiments. The review is closed by a brief discussion of the implications of these measurements on various extensions of the Standard Model of particle physics in Sec. 4.

2. Theory expectation of the branching fractions

In this section, the calculation of the branching fraction of $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ is first discussed in a model independent way, followed by the numerical SM

prediction. For simplicity of notation, the expressions are given solely for the decay $B_s^0 \rightarrow \mu^+ \mu^-$. The corresponding expression for $B^0 \rightarrow \mu^+ \mu^-$ decays can be trivially obtained by exchanging the s with a d quark.

2.1. Model independent discussion

The branching fraction of $B_s^0 \rightarrow \mu^+ \mu^-$ can be expressed as low-energy effective Hamiltonian using the operator product expansion (OPE) which allows to separate the long distance contributions to the decay amplitude from the short distance contributions. The former are relegated to non-perturbative hadronic matrix elements whereas the latter are described by perturbatively calculable Wilson coefficients C_k . A detailed discussion of the concept can be found in Ref. 20. The effective Hamiltonian for $b \rightarrow s \ell^+ \ell^-$ transitions is given by

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e^2}{16\pi^2} \sum_i (C_i O_i + C'_i O'_i) + h.c., \quad (1)$$

where G_F is the Fermi constant, V_{tb} and V_{ts} are elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix and C_i are the Wilson coefficients. The four-fermion operators O_i which are relevant for $b \rightarrow s \ell^+ \ell^-$ decays are:

$$O_{10} = (\bar{s} \gamma_\mu P_L b) (\bar{l} \gamma^\mu \gamma_5 l), \quad (2)$$

$$O_S = m_b (\bar{s} P_R b) (\bar{l} l), \quad (3)$$

$$O_P = m_b (\bar{s} P_R b) (\bar{l} \gamma_5 l), \quad (4)$$

where $P_{L,R} = \frac{1}{2}(1 \mp \gamma_5)$. The corresponding operators O'_i are obtained from the O_i operators by replacing P_L with P_R . The notation used here follows the one of Ref. 21 and 22. The branching ratio of $B_s^0 \rightarrow \mu^+ \mu^-$ decays can then be written in a model independent way as

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = \frac{4G_F^2 \alpha^2}{64\pi^2} f_{B_s}^2 m_{B_s}^3 \tau_{B_s} |V_{tb} V_{ts}^*| \sqrt{1 - \frac{4m_\mu^2}{m_{B_s}^2}} \quad (5)$$

$$\times \left\{ \left(1 - \frac{4m_\mu^2}{m_{B_s}^2} |C_S - C'_S|^2 + |(C_P - C'_P) + 2 \frac{m_\mu}{m_{B_s}} (C_{10} - C'_{10})|^2 \right) \right\},$$

where f_{B_s} is the B_s^0 decay constant, m_{B_s} and τ_{B_s} are the B_s^0 mass and lifetime respectively. The contributions of the scalar (C_S , C'_S) and pseudoscalar (C_P , C'_P) operators enter in the branching fraction without suppression. As the corresponding Wilson coefficients are still largely unconstrained, there is significant room for contributions from New Physics models. In the SM, however, they are predicted to be strictly vanishing. The contribution from electroweak penguin diagrams, contained in the coefficients C_{10} and C'_{10} , is suppressed by a helicity factor $(m_\mu/m_{B_s})^2 \approx 4 \times 10^{-4}$. In the Standard Model, only C_{10} is non-zero and its value is given by the real coefficient C_{10}^{SM} . It is dominated by a Z^0 penguin loop (75%) and a box diagram (24%)²³. The two dominant SM Feynman diagrams contributing to

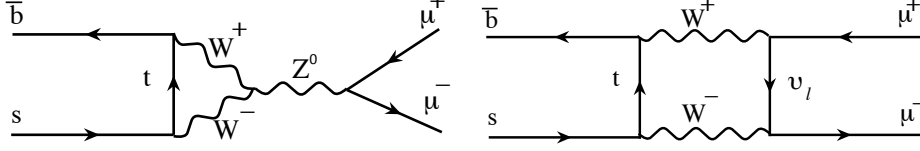


Fig. 1. The two dominating SM Feynman diagrams contributing to the decay $B_s^0 \rightarrow \mu^+ \mu^-$.

the $B_{s,d}^0 \rightarrow \mu^+ \mu^-$ decay are given in Fig. 1. The sensitivity of this decay to scalar and pseudo-scalar interactions enables stringent tests of New Physics models with an extended (pseudo-) scalar sector.

2.2. Standard Model prediction

The numerical value of the Standard Model prediction of the branching fraction of $B_s^0 \rightarrow \mu^+ \mu^-$, as given in Eq. (5), can be calculated as

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = 3.1 \pm 0.2 \times 10^{-9}, \quad (6)$$

where the input parameters as discussed in Ref. 17 are used. The dominant uncertainty comes from the lattice based calculation of the B_s^0 decay constant, which is used here from a recent lattice average²⁴ as $f_{B_s} = 227.7 \pm 6.2$ MeV. The lattice calculations made significant progress recently and thus made this very high precision possible^{25,26,27}. Depending on the choice of the numerical values of the input parameters for both the lattice and the experimental quantities, different authors calculate slightly different branching fractions^{22,23,28,29}, up to 3.6×10^{-9} . A new lattice average by the FLAG-2 group, which represents all big lattice collaborations, is expected by the end of 2012, which will hopefully help to settle these discrepancies in the predicted numerical value of $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$.

Alternatively to the evaluation of Eq. (5), the CKM dependencies and hadronic uncertainties can be eliminated by expressing the branching fraction as a function of the $B_s^0 - \bar{B}_s^0$ meson mass difference ΔM_s . The dependence on the decay constant f_{B_s} is exchanged with the bag parameter \hat{B}_{B_s} , which can be calculated with smaller uncertainties. The branching fraction of $B_s^0 \rightarrow \mu^+ \mu^-$ decays is then given as^{25,30}

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = 4.36 \times 10^{-10} \times \frac{\tau_{B_s^0}}{\hat{B}_{B_s}} \frac{Y^2(v)}{S(v)} \Delta M_s \quad (7)$$

$$= 3.2 \pm 0.2 \times 10^{-9}, \quad (8)$$

where $Y^2(v)$ and $S(v)$ are the two master functions of CMFV models³¹, evaluated in the SM with $v = m_t^2/M_W^2$. In Eq. (8), the numerical value of $\hat{B}_{B_s} = 1.33 \pm 0.06$ from Ref. 24 and 32 is used. It is a remarkable success of the lattice calculations that both values determined in Eq. (6) and Eq. (8) are in excellent agreement.

2.3. Comparison of the calculated BR with experiment

The LHCb collaboration recently measured the width difference in the B_s^0 system in a time dependent analysis of $B_s^0 \rightarrow J/\psi \phi$ decays³³ as $\Delta\Gamma_s = 0.116 \pm 0.019 \text{ ps}^{-1}$. If the calculated value of $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ is compared to a measurement, the finite width difference has to be taken into account as the measured branching fraction is the time integrated one,

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)^{\text{exp}, \langle t \rangle} = \frac{1}{2} \int_0^\infty \langle \Gamma(B_s^0(t) \rightarrow \mu^+ \mu^-) \rangle dt, \quad (9)$$

whereas the the CP averaged branching fraction is calculated. To compare the measured value with a theory prediction, one of the two values has to be corrected, as pointed out recently^{34,35}. Here, a correction of the theoretical value (Eq. (6)) is chosen which reads

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)^{\text{TH}, \langle t \rangle} = \frac{1}{1 - y_s} \times \mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)^{CP} = 3.4 \pm 0.2 \times 10^{-9}, \quad (10)$$

where the CP averaged branching ratio as given in Eq. (6) is taken and y_s is given as

$$y_s = \frac{\Delta\Gamma_s}{2\Gamma_s} = 0.088 \pm 0.014. \quad (11)$$

To compare the SM prediction to a measured branching fraction or an exclusion limit, the value of the branching fraction of $B_s^0 \rightarrow \mu^+ \mu^-$ given in Eq. (10) has to be used.

3. Experimental situation

The search for $B_{s,d}^0 \rightarrow \mu^+ \mu^-$ has been performed both at the B-factories BaBar and Belle as well as at the experiments at the Tevatron and at the LHC. At the B-factories, running at the $\Upsilon(4S)$ resonance, only B^0 decays are accessible whereas at the Tevatron and the LHC both B^0 and B_s^0 mesons are produced.

A comparison of the the $b\bar{b}$ -production cross section at the different experiments is given in Tab. 1. The efficiency of the acceptance and p_T requirements is not corrected for, as the applied selections are typical for analyses at the given experiments. The production cross section at hadron colliders is, depending on collision energy and detector acceptance, between $6 \mu\text{b}$ (CDF, D0) and $94 \mu\text{b}$ (LHCb). Note that the LHCb experiment is a single arm, forward spectrometer which integrates the peak of $b\bar{b}$ -pairs produced in forward direction whereas the general purpose detectors ATLAS and CMS (GPD) have a limited acceptance in the forward or backward region but integrate over the complete remaining phase space. The effect of these acceptances is that the $b\bar{b}$ -production cross section is similar for the LHCb experiment and the GPD. The sensitivity of recent HEP experiments for the searches for $B_{s,d}^0 \rightarrow \mu^+ \mu^-$ can be summarized as follows:

- Compared to the cross sections at the Tevatron and the LHC, the $b\bar{b}$ -production rate at e^+e^- colliders at $\Upsilon(4S)$ is more than three orders of magnitude smaller. Hence, the number of recorded $b\bar{b}$ -pairs is much lower and the searches for $B_{s,d}^0 \rightarrow \mu^+\mu^-$ are with the existing datasets of 500 and 1000 fb^{-1} not competitive. Therefore, the discussion in this review focuses on the hadron collider experiments.
- The Tevatron collider experiments recorded a dataset corresponding to an integrated luminosity of about 10 fb^{-1} in Run II between 2001 and 2011, with an estimated number of 6×10^{10} $b\bar{b}$ -pairs produced per experiment.
- Since the LHC began collecting data in 2010, its experiments have taken over the lead of the field due to the very high $b\bar{b}$ -production cross section. At the general purpose detectors at the LHC, about 4×10^{11} $b\bar{b}$ -pairs have been produced in their dataset of 5 fb^{-1} which has been collected at instantaneous luminosities up to $3.5 \times 10^{33} \text{ s}^{-1} \text{ m}^{-2}$. The LHCb experiment, in contrast to the GPD, leveled its luminosity in 2011 to a constant value of $2 - 3 \times 10^{32} \text{ s}^{-1} \text{ m}^{-2}$, giving a total dataset until the end of 2011 of 1 fb^{-1} which contains 9×10^{10} $b\bar{b}$ -pairs.

Thanks to the very large number of $b\bar{b}$ -pairs produced and the clean experimental signature of the $B_{s,d}^0 \rightarrow \mu^+\mu^-$ decays, the hadron collider experiments have performed the most sensitive searches for the decays $B_{s,d}^0 \rightarrow \mu^+\mu^-$. These are discussed in detail in the remainder of this section.

3.1. Overview of the analyses and search strategies

The search for the rare decays $B_{s,d}^0 \rightarrow \mu^+\mu^-$ has been performed at both Tevatron detectors. The analyses discussed here contain the published CDF analysis using 7 fb^{-1} of $p\bar{p}$ collision data⁴¹ at $\sqrt{s} = 1.96 \text{ TeV}$ as well as their unpublished update⁴² using the full dataset of 10 fb^{-1} , collected until the shutdown of the Tevatron collider. The D0 collaboration published a result⁴³ with a dataset corresponding to an integrated luminosity of 6.1 fb^{-1} . The sensitivity of the D0 detector is about a factor of two worse than the CDF sensitivity as the D0 mass resolution is much

Table 1. $b\bar{b}$ -production cross sections in the acceptance of recent HEP experiments. Data taken from Refs. 36, 37, 38, 39. The cross section for ATLAS, CMS, CDF, LHCb is calculated from a B^+ cross section measurement with $p_T > 5 \text{ GeV}$ (CDF: $p_T > 6 \text{ GeV}$, LHCb $p_T > 0 \text{ GeV}$) using the LHCb measurement of the hadronization fractions⁴⁰. The number of produced $b\bar{b}$ -pairs estimated for the dataset recorded until the end of 2011 and does not include trigger and reconstruction efficiencies.

experiment	acceptance	$\Upsilon(4S)$	$p\bar{p}$, 1.96 TeV	pp , 7 TeV	$b\bar{b}$ -pairs produced
BaBar, Belle	4π	1.05 nb	$6.3 \pm 0.6 \mu\text{b}$	$75 \pm 17 \mu\text{b}$ $94 \pm 8 \mu\text{b}$	$\sim 1 \times 10^9$
CDF	$ \eta < 1$				$\sim 6 \times 10^{10}$
ATLAS, CMS	$ \eta < 2.2$				$\sim 4 \times 10^{11}$
LHCb	$2 < \eta < 6$				$\sim 9 \times 10^{10}$

worse. For the decay $B_s^0 \rightarrow \mu^+\mu^-$, a mass resolution of $24 \text{ MeV}/c^2$ is expected for the CDF experiment whereas $120 \text{ MeV}/c^2$ is expected for the D0 experiment. Due to its lower sensitivity, the D0 result is not discussed further in this review.

At the LHC, the most sensitive measurements of the branching ratios of $B_{s,d}^0 \rightarrow \mu^+\mu^-$ have been made by the LHCb collaboration, which published three measurements^{44,45,46} using 37 pb^{-1} , 370 pb^{-1} and 1 fb^{-1} of $\sqrt{s} = 7 \text{ TeV}$ pp -collision data, respectively. Both general purpose detectors at the LHC have performed searches for $B_{s,d}^0 \rightarrow \mu^+\mu^-$, the ATLAS collaboration⁴⁷ using 2.4 fb^{-1} of data and the CMS collaboration^{48,49}, using 1 fb^{-1} and 5 fb^{-1} of $\sqrt{s} = 7 \text{ TeV}$ pp -collision data.

The invariant mass resolution of the three LHC experiments, estimated for $B_s^0 \rightarrow \mu^+\mu^-$ signal candidates, is

$$\begin{aligned} \text{ATLAS} &: 60 \text{ MeV}/c^2 \ (|\eta| < 1) - 110 \text{ MeV}/c^2 \ (|\eta| > 1.5), \\ \text{CMS} &: 37 \text{ MeV}/c^2 \ (|\eta| \sim 0) - 77 \text{ MeV}/c^2 \ (|\eta| > 1.8), \\ \text{LHCb} &: 24.8 \text{ MeV}/c^2, \end{aligned}$$

which shows the invariant mass resolution of the LHCb forward spectrometer to be on average a factor two better than the CMS resolution and a factor 3-4 better than the ATLAS resolution. This, together with the better separation between signal and background in the boosted region and the cleaner environment compensates for the luminosity, which is a factor 5 lower in LHCb compared to the GPDs.

The measurements performed by the CDF and LHCb collaborations use a similar strategy to search for $B_{s,d}^0 \rightarrow \mu^+\mu^-$ decays. Triggered opposite sign dimuon candidates are further selected to clean up the sample. To maximize the efficiency, a significant amount of background is allowed in the selection. The remaining events are classified in a two-dimensional plane spanned by the invariant mass and the signal likelihood formed from the two-prong decay signature.

In contrast to this, the ATLAS and CMS collaborations use a tight selection which is cut based at CMS and based on a boosted decision tree in ATLAS. Both measurements are binned in rapidity, separating regions with different detector performance and therefore different signal to background ratios.

The event selection and classification with a signal likelihood is discussed in Sec. 3.2. All experiments use a relative normalization to interpret the observed pattern of events as branching ratio, discussed in Sec. 3.3, and then use the modified frequentist method (CL_s) to extract the limits on the branching fractions, as discussed in Sections 3.4 and 3.5. In Sec. 3.6, a combination of the searches for $B_{s,d}^0 \rightarrow \mu^+\mu^-$ of the three LHC experiments is discussed.

3.2. Event selection and Signal Likelihood

All experiments have highly efficient muon triggers, which are used both to select events for the signal and for the normalization channel (see, *e.g.*, Refs. 50 or 51 for a detailed discussion). The dimuon triggers for the $B_s^0 \rightarrow \mu^+\mu^-$ analysis obtained

a high priority in the physics programs of the GPDs, therefore, the high efficiency could be preserved during the high luminosity running.

The fully reconstructed signal candidates can be identified by their clearly separated secondary vertex, exploiting the long B_s^0 meson lifetime. The candidate momentum vector will be aligned with the separation between primary and secondary vertex (pointing), and the two muon candidates will be isolated from other tracks in the event, due to the hard B fragmentation.

Background events tend to be partially reconstructed and shorter lived than the signal. They also have a softer p_T spectrum, lower degree of isolation and a less precise pointing. The dominant combinatorial background consists of sequential semileptonic decays ($b \rightarrow c\mu^- \rightarrow \mu^+\mu^-X$) and double semi-leptonic decays ($b \rightarrow \mu^+X$, $b \rightarrow \mu^-X$). The former mostly populate the lower mass sidebands whereas the latter cover the whole mass range. In the analyses, these combinatorial backgrounds are estimated by an extrapolation from the invariant mass sidebands.

Peaking backgrounds, dominantly originating from misidentified $B_s^0 \rightarrow K^+K^-$, $B^0 \rightarrow \pi^+\pi^-$ and $B^0 \rightarrow K^+\pi^-$ decays, need to be evaluated separately. Here, the invariant mass line shape can be extracted from simulated samples of doubly misidentified $B_{(s)}^0 \rightarrow h^+h'^-$ events. The misidentification rates $K^\pm \rightarrow \mu^\pm$ and $\pi^\pm \rightarrow \mu^\pm$ are extracted in data from control channels such as $D^0 \rightarrow K^-\pi^+$. The contributions of exclusive decays such as $B_c^+ \rightarrow J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$ or $B_s^0 \rightarrow \mu^+\mu^-\gamma$ events is found to be negligible.

3.2.1. *Classification of events*

The CDF and LHCb measurements use a loose selection that allows significant amounts of background and then use a multivariate selection to discriminate signal from background. In this review only the more sensitive analysis of the LHCb collaboration is discussed in detail.

LHCb analysis of 1fb^{-1}

The selected sample is first cleaned with a multivariate classifier based on six variables that can be used equally for the signal and the normalization channels. This selection removes 80% of the residual background, while retaining 92% of the signal. Applying it improves the performance of the main multivariate selection described below.

About 17 000 dimuon candidates pass this selection while 11.6 $B_s^0 \rightarrow \mu^+\mu^-$ and 1.3 $B^0 \rightarrow \mu^+\mu^-$ candidates are expected, assuming SM rates. The selected dimuon candidates are classified in a binned two-dimensional space formed by the dimuon invariant mass and the output of another boosted decision tree (BDT) which combines nine variables to optimally exploit geometrical and kinematic information. The variables used to construct the BDT include the B candidate pointing, decay

time and p_T , the displacement of the muons from the primary vertex and the distance between the muon pairs as well as a measure of the track and B candidate isolation.

No data were used for the choice of the variables and the subsequent training of the BDT, to avoid biasing the results. Instead the BDT was trained using simulated samples, $B_s^0 \rightarrow \mu^+ \mu^-$ for signal and $b\bar{b} \rightarrow \mu^+ \mu^- X$ for background. It is defined such that for the signal it is approximately uniformly distributed between zero and one, while for the background it peaks at zero. Of the 17 321 dimuon candidates after the selection^a, only 95 are in the region of high signal likelihood ($\text{BDT} > 0.5$) which contains about 50% of the expected signal.

The probability for a signal event to have a given BDT value is obtained from data using an inclusive $B_{(s)}^0 \rightarrow h^+ h'^-$ sample. Only events triggered independently of the signal candidates are considered. The fraction of $B_{(s)}^0 \rightarrow h^+ h'^-$ signal candidates in each bin of BDT is determined by fitting the $h^+ h'^-$ invariant mass distribution.

The chosen number and size of the bins are a compromise between maximizing the number of bins and the necessity to have enough $B_{(s)}^0 \rightarrow h^+ h'^-$ events to calibrate the signal BDT and enough background in the mass sidebands to estimate the combinatorial background in the B_s^0 and B^0 mass regions. The BDT range is thus divided into eight bins and the invariant mass range into nine bins.

CDF analysis

The CDF collaboration uses a similar multivariate classifier (called ν_N) which is based on an artificial neural net combining 14 input variables. The sample is binned in 8 bins of ν_N and 5 bins of invariant mass. The signal shape of ν_N is taken from the simulation and cross checked with $B^- \rightarrow J/\psi K^-$ events.

The evaluation of the compatibility of the observed pattern of events with the background or signal hypotheses is then performed in these bins and will be described in detail in Sec. 3.4 and 3.5.

3.2.2. Cut based separation of signal and background

The ATLAS and CMS measurements use a tight selection of the signal on which the background is cut away. The ATLAS selection is based on a boosted decision tree combining 14 variables. To ensure that the data are reproduced, the simulation is tuned with an iterative reweighting procedure.

The CMS collaboration optimizes a cut based selection on MC signal and data sideband events. A multivariate analysis is in preparation which is expected to give an improvement of up to 20% relative to the cut based analysis.

^aThe number of events quoted here are in a mass range between $4900 \text{ MeV}/c^2$ and $6000 \text{ MeV}/c^2$.

3.3. Relative Normalization

The branching fraction for the $B_{(s)}^0 \rightarrow \mu^+\mu^-$ signal is measured by all analyses discussed here relative to a channel with a well known branching fraction as

$$\mathcal{B} = \mathcal{B}_{\text{norm}} \times \frac{\epsilon_{\text{norm}}}{\epsilon_{\text{sig}}} \times \frac{f_{\text{norm}}}{f_{d(s)}} \times \frac{N_{B_{(s)}^0 \rightarrow \mu^+\mu^-}}{N_{\text{norm}}}, \quad (12)$$

where the branching fraction is determined by $\mathcal{B}_{\text{norm}}$, the branching fraction of the normalization channel. It needs to be corrected by the relative efficiencies $\epsilon_{\text{norm}}/\epsilon_{\text{sig}}$, the probabilities that a b quark fragments into a $B_{(s)}^0$ and into the b hadron involved for the chosen normalization mode $f_{\text{norm}}/f_{d(s)}$ and by the ratio of observed candidates in the signal and the normalization mode, $N_{B_{(s)}^0 \rightarrow \mu^+\mu^-}/N_{\text{norm}}$. This relative normalization has many advantages over a absolute normalization: the luminosity and production cross section is not required to be known and the determination of relative efficiencies is more robust than the determination of absolute efficiencies.

All experiments discussed here measure $\mathcal{B}(B_{(s)}^0 \rightarrow \mu^+\mu^-)$ relative to the $B^- \rightarrow J/\psi(\mu^+\mu^-)K^-$ channel, which has a similar muon identification and trigger, which minimizes systematic uncertainties. Its branching fraction is known to a precision of 3.5%⁵². However, as it is a three body decay, the detector acceptance as well as the kinematic distributions of the tracks differ and need to be corrected for.

The LHCb collaboration performs the only measurement that uses the weighted average of three normalization channels, additionally $B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ and $B_{(s)}^0 \rightarrow h^+h'^-$. These three channels have different advantages as normalization: the $B_s^0 \rightarrow J/\psi\phi$ decay is a B_s^0 decay and hence the fraction of hadronization probabilities cancels. However, the branching fraction of this decay is currently known to a precision of only 26%⁵³. The $B_{(s)}^0 \rightarrow h^+h'^-$ channel has an identical two body signature as the signal decay, but has been selected by very different hadronic trigger than the dimuon signal. Therefore, only $B_{(s)}^0 \rightarrow h^+h'^-$ candidates that are unbiased by the trigger selections are used for the normalization. While the normalization of the average of three channels significantly improves the robustness of the analysis and reduces the dependence on the simulation, the B^+ channel strongly dominates the average.

The mass distributions of the selected $B^- \rightarrow J/\psi K^-$ candidates are shown in Fig. 2 for the four experiment discussed here. Please note the different scale of the figures, while the resolution of the LHCb and CDF experiments is good enough to clearly separate the shoulder of partially reconstructed $B_d^0 \rightarrow J/\psi K^{*0}$ events from the signal, the ATLAS and CMS experiments need to model it.

The relative efficiencies are determined by all experiments from the simulation and have been verified with $B_s^0 \rightarrow J/\psi\phi$, $B_d^0 \rightarrow J/\psi K^{*0}$ and $B_{(s)}^0 \rightarrow h^+h'^-$ events in data. The uncertainties the efficiency ratio contributes to the normalization factors vary between 3% (LHCb) and 16% (ATLAS end-cap).

For the normalization to B^+ candidates, the ratio of b -hadronization fractions need to be included in Eq. 12. The CDF experiment uses the PDG average for the

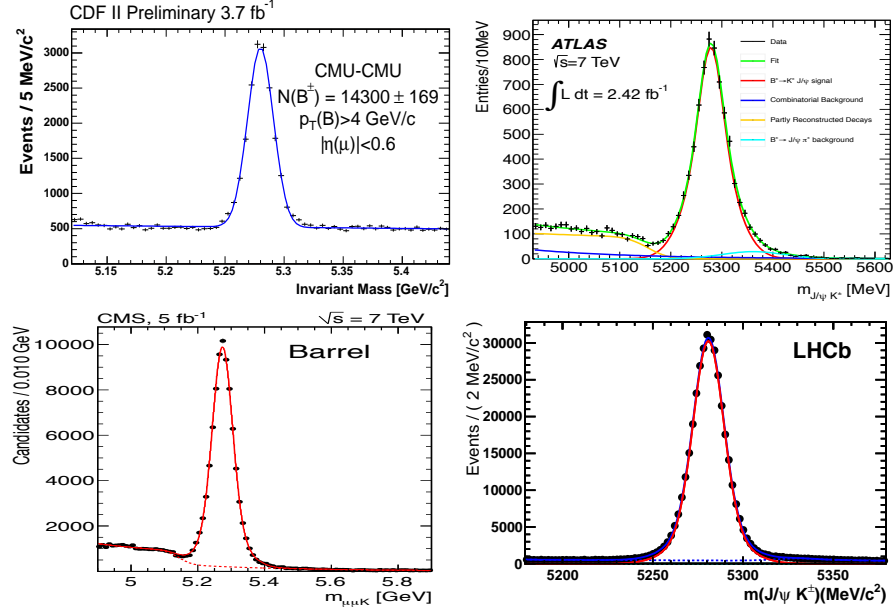


Fig. 2. Invariant mass distribution of the $B^- \rightarrow J/\psi K^-$ candidates used in the normalization procedure. Note the different scale on the Figures. Reproduced from CDF⁴², ATLAS⁴⁷, CMS⁴⁹ and LHCb⁴⁶.

hadronization fraction⁵², which has an uncertainty of 10%. The LHC experiments all use the measurement performed by the LHCb collaboration⁴⁰, which has an uncertainty of 7.8%. The ATLAS and CMS collaborations use this measurement assuming it is valid for all momentum and rapidity regions.

3.4. Extraction of the branching ratio

The branching fraction limit is extracted from the pattern of events using the modified frequentist method^{54,55} (CL_s), where the background only (B) and signal plus background (S+B) hypotheses are tested. The logarithmic form of the likelihood ratio

$$-2 \ln Q = -2 \ln \frac{\mathcal{L}_{S+B}}{\mathcal{L}_B} = -2 \ln \prod_i \frac{e^{s_i+b_i} (s_i + b_i)^{d_i} / d_i!}{e^{-b_i} (b_i)^{d_i} / d_i!}, \quad (13)$$

depends on the expected number of signal (s_i) and background (b_i) events in each bin i , as well as on the observed number of events (d_i). The method provides a measure of the compatibility of the observed distribution with the signal plus background hypothesis, CL_{s+b} , as well as a measure for the compatibility with the background only hypothesis, CL_b , which are both derived from pseudo experiments. The fraction of signal plus background (background only) events with a value of $-2 \ln Q$ smaller than or equal to the observed value is computed. Systematic uncertainties

are included in the pseudo experiments as nuisance parameters using Poisson distributions for the uncertainties with statistical nature and bifurcated Gaussians for the others.

The ratio of both confidence levels, $\text{CL}_s = \text{CL}_{s+b}/\text{CL}_b$, is used to set an exclusion limit on the branching fraction of the two decays to protect against negative statistical fluctuations of the background which could lead to an exclusion of the null hypothesis without experimental sensitivity. To claim evidence of the decays, $1 - \text{CL}_b$ is used as a p-value. A 95% confidence level exclusion corresponds to a CL_s value of 0.05 and a 3σ evidence to a value of $1 - \text{CL}_b = 2.7 \times 10^{-3}$ for a one sided definition of the significance.

3.5. Results

The CDF collaboration observes an excess of candidates in the $B_s^0 \rightarrow \mu^+ \mu^-$ signal region. It is concentrated in bins with a high signal likelihood, $\nu_N > 0.97$. There is also an excess in the medium likelihood bins, $0.97 < \nu_N < 0.987$, which appears to be a statistical fluctuation of the background as there is no significant expectation of $B_s^0 \rightarrow \mu^+ \mu^-$ signal consistent with the observation in the two highest bins. The p-value for a background-only hypothesis is 0.94%, when considering only the two highest ν_N bins the p-value becomes 2.1%. The branching fraction corresponding to the observed pattern of events is determined from a log-likelihood fit and is determined to be

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)^{\text{CDF}} = 13_{-7}^{+9} \times 10^{-9}. \quad (14)$$

Additionally, bounds are set on the branching fraction using the CL_s method as $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) < 31 \times 10^{-9}$ at 95% CL.

The excess seen by the CDF collaboration has not been confirmed by the LHC experiments, which are summarized in Tab. 2. The ATLAS collaboration analyses 2.4 fb^{-1} of data and observes no excess of candidates and determines an upper limit of $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) < 22 \times 10^{-9}$, being very close to the value expected from background extrapolation.

The CMS collaboration analyses 4.9 fb^{-1} of data and observes the pattern of events shown in Fig. 3 (top). Six events are observed in the $B_s^0 \rightarrow \mu^+ \mu^-$ signal region, of which two are in the barrel and four in the end-cap. This pattern of events is consistent with the expectation from the Standard Model branching fraction plus background. The upper limits combined for both regions are at 95% CL:

$$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)^{\text{CMS}} < 7.7 \times 10^{-9} \text{ and} \quad (15)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-)^{\text{CMS}} < 1.8 \times 10^{-9}, \quad (16)$$

which is well comparable to the median expected upper limits of 8.4×10^{-9} for the B_s^0 mode and 1.6×10^{-9} for B^0 mode.

The LHCb collaboration analyses 1 fb^{-1} of data and observes the distribution of the invariant mass for $\text{BDT} > 0.5$ as shown in Fig. 3 (bottom) for B_s^0 and B^0

Table 2. Expected and observed upper limit on $\mathcal{B}(B_{s,d}^0 \rightarrow \mu^+\mu^-)$ for the most recent measurements, all given at 95% CL. For the expected limits, it is indicated if the SM signal plus background or the background only case is assumed. The CDF collaboration observes an excess of $B_s^0 \rightarrow \mu^+\mu^-$ candidates and quotes bounds at 95% CL on the branching fraction: $0.8 \times 10^{-9} < \mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) < 34 \times 10^{-9}$.

experiment	luminosity	$B_s^0 \rightarrow \mu^+\mu^-$		$B^0 \rightarrow \mu^+\mu^-$	
		expected	observed $\times 10^{-9}$	expected	observed
D0	6 fb^{-1}	40 (bkg)	51	-	-
CDF	10 fb^{-1}	13 (bkg)	31	40 (bkg)	46
ATLAS	2.4 fb^{-1}	23 (bkg)	22	-	-
CMS	4.9 fb^{-1}	8.4 (SM+bkg)	7.7	1.6 (bkg)	1.8
LHCb	1 fb^{-1}	7.2 (SM+bkg)	4.5	1.1 (bkg)	1.0

decays. The expected limits are computed allowing the presence of $B_s^0 \rightarrow \mu^+\mu^-$ events according to the SM rate, including cross feed between the two modes. A limit of $7.2(1.1) \times 10^{-9}$ is expected for the B_s^0 and B^0 modes respectively. The analysis of the observed pattern of events gives

$$\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)^{\text{LHCb}} < 4.5 \times 10^{-9} \text{ and} \quad (17)$$

$$\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)^{\text{LHCb}} < 1.0 \times 10^{-9}. \quad (18)$$

This is a downward fluctuation of 1σ for the B_s^0 mode with respect to the SM expectation. A simultaneous unbinned likelihood fit to the mass projections in the eight BDT bins has been performed to determine the $B_s^0 \rightarrow \mu^+\mu^-$ branching fraction. The fit gives

$$\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)^{\text{LHCb}} = 0.8_{-1.3}^{+1.8} \times 10^{-9}, \quad (19)$$

where the central value is extracted from the maximum of the logarithm of the profile likelihood and the uncertainty contains both statistical fluctuations and systematic uncertainties.

In comparison to the CMS measurement, LHCb has a 15% higher sensitivity which shows that the better detector resolution (both in mass and vertex precision) and the more advanced analysis overbalance the factor five higher luminosity that the CMS experiment has recorded. Due to a downward fluctuation of the data, the measured upper exclusion limit on the decay $B_s^0 \rightarrow \mu^+\mu^-$ is about 40% below the CMS limit.

3.6. Combination of $B_{s,d}^0 \rightarrow \mu^+\mu^-$ measurements

The results of the analyses performed on the 2011 LHC dataset of the ATLAS, CMS and LHCb collaborations have been combined⁵⁶. As the LHCb collaboration has also published a measurement using the data collected in 2010, the measurements performed on both datasets are used in the combination.

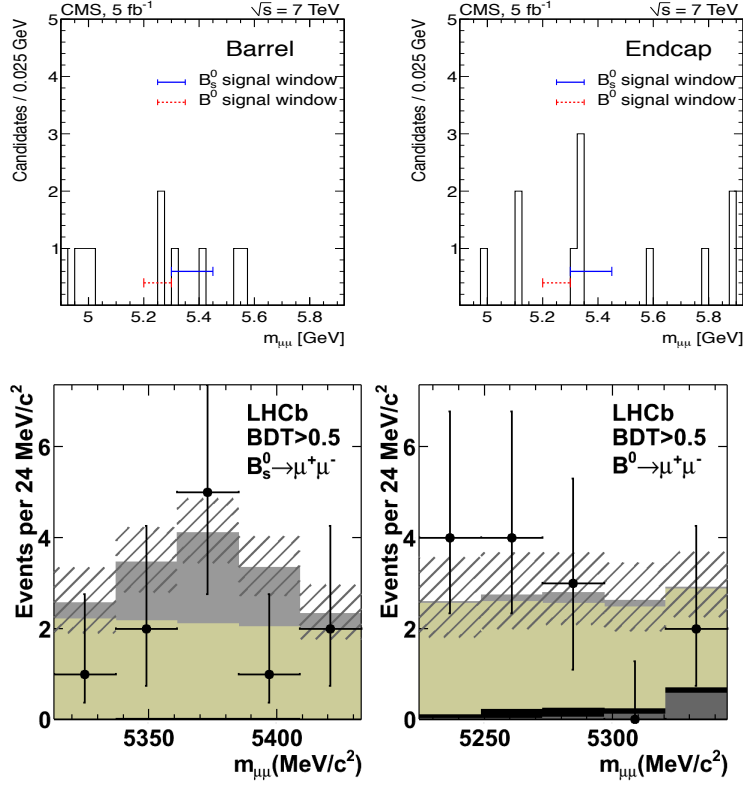


Fig. 3. Distribution of selected signal candidates. Top: Events observed in CMS in the barrel (left) and end-cap (right) channels. Bottom: Events observed in LHCb in the B_s^0 channel (left) and the B^0 channel (right) for $\text{BDT} > 0.5$ and expectation for, from top, SM signal (gray), combinatorial background (light gray), $B_{(s)}^0 \rightarrow h^+h^-$ background (black) and cross feed between both modes (dark gray). The hatched area depicts the uncertainty on the total background expectation. Figures reproduced from CMS⁴⁹ and LHCb⁴⁶.

The combination uses the modified frequentist approach, as described in Sec. 3.4, to combine the individual measurements. The expected and observed CL_s values of the three experiments are shown in Fig. 4 (left) as a function of the assumed branching ratio of $B_s^0 \rightarrow \mu^+\mu^-$. A exclusion limit of $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-) < 4.2 \times 10^{-9}$ is determined. The compatibility of the determined branching ratio with the Standard Model value (see Sec. 2.3) is evaluated as $1 - \text{CL}_{s+b} = 0.84$. The combination of the three LHC experiments shows a moderate excess of the branching ratio of $B_s^0 \rightarrow \mu^+\mu^-$ over the background only hypothesis of $1 - \text{CL}_b = 0.05$. The combined CL_s curve for the $B^0 \rightarrow \mu^+\mu^-$ decay is shown in Fig. 4 (right). The observed upper limit on the branching fraction is found to be 8.1×10^{-10} at 95% CL.

The combined upper limits on $B_s^0 \rightarrow \mu^+\mu^-$ and $B^0 \rightarrow \mu^+\mu^-$ are summarized in Tab. 3, they improve the limits obtained by the individual experiments and represent the best existing limits on these decays. The combination improves the expected limit with respect to the LHCb analysis alone by 15% assuming a signal

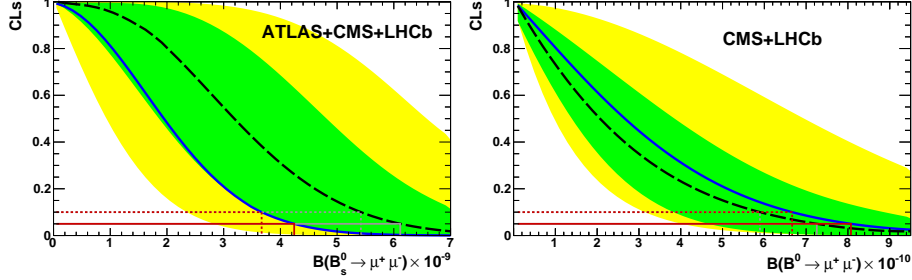


Fig. 4. CL_s as a function of the assumed \mathcal{B} for $B_s^0 \rightarrow \mu^+ \mu^-$ (left) and $B^0 \rightarrow \mu^+ \mu^-$ (right). The dashed black curves are the medians of the expected CL_s distributions, if background and SM signal were observed (for $B_s^0 \rightarrow \mu^+ \mu^-$) and in absence of signal ($B^0 \rightarrow \mu^+ \mu^-$). The green (yellow) areas cover the $\pm 1(2)\sigma$ intervals. The solid blue curves are the observed CL_s . The upper limits at 90 % (95 %) C.L. are indicated by the dotted (solid) horizontal lines. Figure reproduced from Ref. 56.

Table 3. Expected and observed upper limit on $\mathcal{B}(B_{s,d}^0 \rightarrow \mu^+ \mu^-)$ for the combination of the LHC analyses, all given at 95% CL. The expected limits correspond to the median cases in which only background or SM signal and background events were observed.

	Expected		Observed
	bkg only	SM + bkg	
$\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) \times 10^{-9}$	2.3	6.1	4.2
$\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) \times 10^{-10}$	7.3		8.1

at the SM rate any by 32% assuming the absence of any signal. Despite this largely improved sensitivity, the effect in the observed upper limit is reduced as the LHCb experiment observes a downward fluctuation of signal events with respect to the SM prediction.

4. Implications

The exclusion limit on the branching fraction of the decay $B_s^0 \rightarrow \mu^+ \mu^-$ only 20% above the SM expectation provides stringent tests of possible extensions of the Standard Model. Especially the scalar sector, as discussed in Sec. 2, is strongly constrained by the exclusion of large enhancements in the branching fraction.

The implications of the existing measurements are discussed in the context of two simplified variants of the minimal supersymmetric extension of the Standard Model (MSSM)^{57,58}: the constrained MSSM (CMSSM)⁵⁹ and non-universal Higgs Masses of type 1 (NUHM1)⁶⁰. In the CMSSM, the number of free parameters is reduced to five: $m_0, m_{1/2}$ and A_0 , denoting common scalar, fermionic and trilinear soft supersymmetry-breaking parameters at the GUT scale, and $\tan \beta, \text{sgn } \mu$ denoting the ratio of the vacuum expectation values of the two Higgs fields and the sign of the mass term respectively. In the NUHM1 model, the universality condition for the Higgs bosons are decoupled from the other scalars, adding two extra parameters to the model.

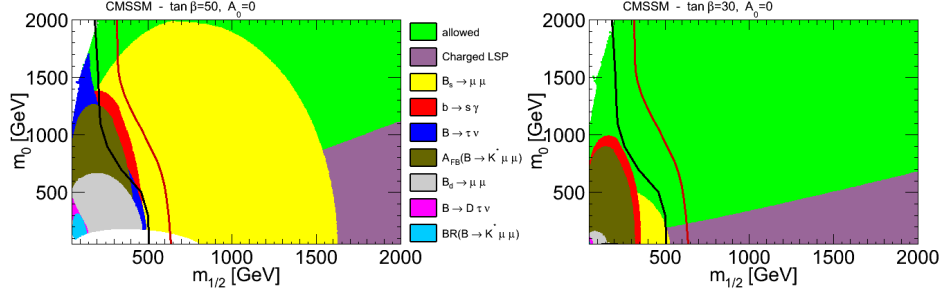


Fig. 5. Constraints from flavor observables in CMSSM in the plane $(m_{1/2}, m_0)$ with $A_0 = 0$, for $\tan\beta = 50$ in the left and 30 in the right. The black line corresponds to the CMS exclusion limit with 1.1 fb^{-1} of data and the red line to the CMS exclusion limit with 4.4 fb^{-1} of data. Figures reproduced from Ref. 61.

A analysis of the constraints imposed by the different flavor observables in the CMSSM model is done, for example, in Refs. 23 or 61. Fig. 5 shows the plane $(m_{1/2}, m_0)$ for large (left) and moderate (right) values of $\tan\beta$ in the CMSSM. Direct search limits from CMS are superimposed for comparison. It can be seen that, at large values of $\tan\beta$, the constraints from $B_s^0 \rightarrow \mu^+ \mu^-$ are stronger than those from direct searches. At smaller values of $\tan\beta$, the flavor observables start to lose importance compared to direct searches. In this regime the other flavor observables, in particular the observables measured in $B^0 \rightarrow K^* \mu^+ \mu^-$ decays, loose less sensitivity and hence play a complementary role. In more general MSSM models, the parameter space is significantly less constrained.

The influence of the searches for $B_{s,d}^0 \rightarrow \mu^+ \mu^-$ on various supersymmetric models can be seen in a global analysis of the constraints on these models from the measurements in the high p_T and in the flavor sector⁶². A global fit to all input excluding $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$ predicts in the CMSSM an enhancement up to a factor of two and in NUHM1 up to factor of three. The current limits on $B_s^0 \rightarrow \mu^+ \mu^-$ exclude the minimum of the global fits in both the CMSSM and NUHM1 case.

Another aspect is the interplay between searches for $B_{s,d}^0 \rightarrow \mu^+ \mu^-$ and Higgs physics, since any viable model point has to be in agreement with all direct and indirect measurements. For example, the Higgs boson recently discovered^{63,64} at a mass of about $125 \text{ GeV}/c^2$ rules out MSSM scenarios in which the signal corresponds to the heaviest CP-even Higgs (as opposed to the lightest Higgs), as this would imply a too light pseudo-scalar Higgs to be consistent with the existing constraints on $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)$. See Ref. 65 for a more detailed discussion.

It is clear that with more precise measurements on the branching fraction of $B_{s,d}^0 \rightarrow \mu^+ \mu^-$, large parts of the supersymmetric parameter space could be disfavoured, in particular the region of high values of $\tan\beta$. Because of the difficulty to access this range in direct searches, the measurements of $B_{s,d}^0 \rightarrow \mu^+ \mu^-$ are a crucial element in the exploration or exclusion of supersymmetry. Also, a measurement of $\text{BR}(B_s^0 \rightarrow \mu^+ \mu^-)$ lower than the SM prediction would rule out a large variety of

supersymmetric models.

Once an evidence for the decay $B_s^0 \rightarrow \mu^+ \mu^-$ is found, the ratio of the branching fraction of the $B_s^0 \rightarrow \mu^+ \mu^-$ and the $B^0 \rightarrow \mu^+ \mu^-$ decays becomes accessible. This ratio is essentially free of hadronic uncertainties and is predicted to be³⁰

$$\frac{\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-)}{\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-)} = \frac{\hat{B}_{B_d} \tau_{B_s} \Delta M_s}{\hat{B}_{B_s} \tau_{B_d} \Delta M_d} \stackrel{\text{MFV}}{\approx} 32, \quad (20)$$

where the lifetime $\tau_{B_{s,d}}$ and mixing frequency $\Delta M_{s,d}$ are experimental input and ratio of the bag parameters is well known, $\hat{B}_{B_d}/\hat{B}_{B_s} = 0.95 \pm 0.06$. A measurement of this ratio provides a stringent test of the minimal flavor violation (MFV) hypothesis⁶⁶.

The correlation between both decays also allows to discriminate between classes of new physics models, as shown in Fig. 6. A large part of the parameter space of the supersymmetric models, where $\tan \beta$ can be large, is ruled out by the constraints. However, in models where NP enters via the semi-leptonic operators $O_{10}^{(\prime)}$, such as the Standard Model with a sequential fourth generation (SM4) or Randall-Sundrum models (RSc) are starting to be probed only now. See Ref. 67 for a detailed discussion of all models shown in Fig. 6.

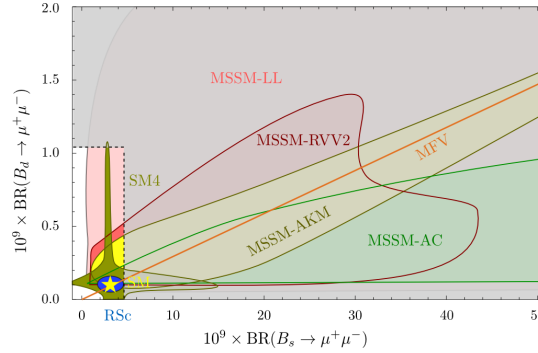


Fig. 6. Correlation between the branching fractions of the $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ decays in MFV, the SM4 and four SUSY flavor models. The SM point is marked by a star and the grey area is excluded by the LHCb measurement given in Eq. (17) and (18). Figure reproduced from Ref. 68.

5. Summary

The experimental situation of the searches for the very rare decays $B_s^0 \rightarrow \mu^+ \mu^-$ and $B^0 \rightarrow \mu^+ \mu^-$ has been reviewed. The data contains a moderate excess of $B_s^0 \rightarrow \mu^+ \mu^-$ signal candidates with a p-value of the background only hypothesis of 0.05. The current most precise upper exclusion limits from a single experiment are measured by the LHCb collaboration as $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) < 4.5 \times 10^{-9}$, the combination of the three LHC experiments is even more stringent at 4.2×10^{-9} . This limit is only

20% above the SM prediction, which puts tight constraints on various extensions of the Standard Model, especially on supersymmetric models at high values of $\tan\beta$. The decay $B^0 \rightarrow \mu^+\mu^-$ is constrained by the existing measurements to an upper exclusion limit of 8×10^{-10} .

The tight experimental bounds on $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$ precludes the optimistic predictions by various supersymmetric models of enhancements of several orders of magnitude. However, a possible exclusion of the Standard Model rate would still be a clear indication of New Physics. It needs also to be noted that the bounds on $\mathcal{B}(B^0 \rightarrow \mu^+\mu^-)$ are still one order of magnitude above the predicted SM value, which means that the most promising channel to search for New Physics could be the B^0 and not the B_s^0 decay.

The CMS and LHCb collaborations have excellent prospects to observe the decay $B_s^0 \rightarrow \mu^+\mu^-$ with the dataset collected in 2012. This observation, and the precision measurement of $\mathcal{B}(B_s^0 \rightarrow \mu^+\mu^-)$ in the coming years will allow to strongly constrain the scalar sector of any extension of the Standard Model. The next step will be to limit and later measure the ratio of the decay rates of $B_s^0 \rightarrow \mu^+\mu^-/B^0 \rightarrow \mu^+\mu^-$, which allows a stringent test of the hypothesis of minimal flavor violation and a good discrimination between various extensions of the Standard Model.

Acknowledgments

I would like to thank many of my colleagues from LHCb for pleasant and fruitful collaboration, specially Frederic Teubert and Diego Martinez Santos for countless discussions about rare decay analyses and Giulia Manca for insightful discussions about production cross sections. Special thanks also to Marco Gersabeck for proof-reading this manuscript.

Furthermore, I acknowledge the support of a Marie Curie Action: Cofunding of the CERN Fellowship Programme (COFUND-CERN) of the European Communitys Seventh Framework Programme under contract number (PCOFUND-GA-2008-229600).

References

1. S. R. Choudhury and N. Gaur *Phys.Lett.* **B451** (1999) 86–92, [arXiv:hep-ph/9810307](#) [hep-ph].
2. J. R. Ellis, K. A. Olive, and V. C. Spanos *Phys.Lett.* **B624** (2005) 47–59, [arXiv:hep-ph/0504196](#) [hep-ph].
3. M. S. Carena, A. Menon, R. Noriega-Papaqui, A. Szykman, and C. Wagner *Phys.Rev.* **D74** (2006) 015009, [arXiv:hep-ph/0603106](#) [hep-ph].
4. J. R. Ellis, S. Heinemeyer, K. Olive, and G. Weiglein *Phys.Lett.* **B653** (2007) 292–299, [arXiv:0706.0977](#) [hep-ph].
5. F. Mahmoudi *JHEP* **0712** (2007) 026, [arXiv:0710.3791](#) [hep-ph].
6. E. Golowich, J. Hewett, S. Pakvasa, A. A. Petrov, and G. K. Yeghiyan *Phys.Rev.* **D83** (2011) 114017, [arXiv:1102.0009](#) [hep-ph].
7. A. Akeroyd, F. Mahmoudi, and D. Martinez Santos *JHEP* **1112** (2011) 088, [arXiv:1108.3018](#) [hep-ph].

8. C.-S. Huang, W. Liao, and Q.-S. Yan *Phys.Rev.* **D59** (1999) 011701, [arXiv:hep-ph/9803460](#) [hep-ph].
9. K. Babu and C. F. Kolda *Phys.Rev.Lett.* **84** (2000) 228–231, [arXiv:hep-ph/9909476](#) [hep-ph].
10. C. Hamzaoui, M. Pospelov, and M. Toharia *Phys.Rev.* **D59** (1999) 095005, [arXiv:hep-ph/9807350](#) [hep-ph].
11. C.-S. Huang, W. Liao, Q.-S. Yan, and S.-H. Zhu *Phys.Rev.* **D63** (2001) 114021, [arXiv:hep-ph/0006250](#) [hep-ph].
12. M. Blanke, A. J. Buras, A. Poschenrieder, S. Recksiegel, C. Tarantino, *et al.* *JHEP* **0701** (2007) 066, [arXiv:hep-ph/0610298](#) [hep-ph].
13. M. Blanke, A. J. Buras, B. Duling, K. Gemmler, and S. Gori *JHEP* **0903** (2009) 108, [arXiv:0812.3803](#) [hep-ph].
14. W. Liu, C.-X. Yue, and H.-D. Yang *Phys.Rev.* **D79** (2009) 034008, [arXiv:0901.3463](#) [hep-ph].
15. M. Bauer, S. Casagrande, U. Haisch, and M. Neubert *JHEP* **1009** (2010) 017, [arXiv:0912.1625](#) [hep-ph].
16. A. J. Buras, B. Duling, T. Feldmann, T. Heidsieck, and C. Promberger *JHEP* **1009** (2010) 104, [arXiv:1006.5356](#) [hep-ph].
17. A. J. Buras and J. Girrbach [arXiv:1204.5064](#) [hep-ph].
18. P. H. Chankowski and J. Rosiek *Acta Phys.Polon.* **B33** (2002) 2329–2354, [arXiv:hep-ph/0207242](#) [hep-ph].
19. R. Arnowitt, B. Dutta, T. Kamon, and M. Tanaka *Physics Letters B* **538** no. 12, (2002) 121 – 129.
20. R. Fleischer *Int.J.Mod.Phys* **A12** (1997) 2459.
21. C. Bobeth, T. Ewerth, F. Kruger, and J. Urban *Phys.Rev.* **D64** (2001) 074014, [arXiv:hep-ph/0104284](#) [hep-ph].
22. W. Altmannshofer, P. Paradisi, and D. M. Straub *JHEP* **1204** (2012) 008, [arXiv:1111.1257](#) [hep-ph].
23. F. Mahmoudi, S. Neshatpour, and J. Orloff [arXiv:1205.1845](#) [hep-ph].
24. J. Laiho, E. Lunghi, and R. S. Van de Water *Phys.Rev.* **D81** (2010) 034503, [arXiv:0910.2928](#) [hep-ph].
25. **HPQCD collaboration**, E. Gamiz, C. T. Davies, G. P. Lepage, J. Shigemitsu, and M. Wingate *Phys.Rev.* **D80** (2009) 014503, [arXiv:0902.1815](#) [hep-lat].
26. H. Na, C. J. Monahan, C. T. Davies, R. Horgan, G. P. Lepage, *et al.* [arXiv:1202.4914](#) [hep-lat].
27. C. Davies *PoS LATTICE2011* (2011) 019, [arXiv:1203.3862](#) [hep-lat].
28. J. Charles, O. Deschamps, S. Descotes-Genon, R. Itoh, H. Lacker, *et al.* *Phys.Rev.* **D84** (2011) 033005, [arXiv:1106.4041](#) [hep-ph].
29. **UTFit collaboration** <http://utfit.org>.
30. A. J. Buras *Phys.Lett.* **B566** (2003) 115–119, [arXiv:hep-ph/0303060](#) [hep-ph].
31. A. J. Buras *Acta Phys.Polon.* **B34** (2003) 5615–5668, [arXiv:hep-ph/0310208](#) [hep-ph].
32. **HPQCD collaboration**, J. Shigemitsu *et al.* *PoS LAT2009* (2009) 251, [arXiv:0910.4131](#) [hep-lat].
33. **LHCb collaboration**. LHCb-CONF-2012-002.
34. K. de Bruyn, R. Fleischer, R. Knegjens, P. Koppenburg, M. Merk, *et al.* [arXiv:1204.1737](#) [hep-ph].
35. K. de Bruyn, R. Fleischer, R. Knegjens, P. Koppenburg, M. Merk, *et al.* [arXiv:1204.1735](#) [hep-ph].
36. **BABAR collaboration**. SLAC Report SLAC-R-504.

37. **CDF collaboration** *Phys. Rev. D* **75** (Jan, 2007) 012010.
38. **LHCb collaboration**, R. Aaij *et al.* *JHEP* **1204** (2012) 093, [arXiv:1202.4812 \[hep-ex\]](#).
39. **CMS collaboration**, S. Chatrchyan *et al.* *Phys.Rev.Lett.* **106** (2011) 252001, [arXiv:1104.2892 \[hep-ex\]](#).
40. **LHCb collaboration**, e. a. Aaij *Phys. Rev. D* **85** (Feb, 2012) 032008.
41. **CDF collaboration**, T. Aaltonen *et al.* *Phys.Rev.Lett.* **107** (2011) 239903, [arXiv:1107.2304 \[hep-ex\]](#). 7 pages, 1 figure/ version accepted by PRL.
42. **CDF collaboration**
<http://www-cdf.fnal.gov/physics/new/bottom/120209.blessed-bmumu10fb/>.
43. **D0 collaboration**, V. M. Abazov *et al.* *Phys.Lett.* **B693** (2010) 539–544, [arXiv:1006.3469 \[hep-ex\]](#).
44. **LHCb collaboration**, R. Aaij *et al.* *Phys.Lett.* **B699** (2011) 330–340, [arXiv:1103.2465 \[hep-ex\]](#).
45. **LHCb collaboration**, R. Aaij *et al.* *Phys.Lett.* **B708** (2012) 55–67, [arXiv:1112.1600 \[hep-ex\]](#).
46. **LHCb collaboration**, R. Aaij *et al.* *Phys. Rev. Lett.* **108**, **231801** (2012) , [arXiv:1203.4493 \[hep-ex\]](#).
47. **ATLAS collaboration**, G. Aad *et al.* *Phys. Lett.* **B713** (2012) 387, [arXiv:1204.0735 \[hep-ex\]](#).
48. **CMS collaboration**, S. Chatrchyan *et al.* *Phys.Rev.Lett.* **107** (2011) 191802, [arXiv:1107.5834 \[hep-ex\]](#).
49. **CMS collaboration**, S. Chatrchyan *et al.* [arXiv:1203.3976 \[hep-ex\]](#).
50. **ATLAS collaboration**, G. Aad *et al.* *Eur.Phys.J.* **C72** (2012) 1849, [arXiv:1110.1530 \[hep-ex\]](#).
51. R. Aaij and J. Albrecht Tech. Rep. LHCb-PUB-2011-017.
CERN-LHCb-PUB-2011-017, CERN, Geneva, Sep, 2011.
52. **Particle Data Group**, K. Nakamura *et al.* *J. Phys. G* **37** (2010) 075021.
53. **Belle collaboration**, R. Louvot [arXiv:0905.4345 \[hep-ex\]](#).
54. T. Junk *Nucl.Instrum.Meth.* **A434** (1999) 435–443, [arXiv:hep-ex/9902006 \[hep-ex\]](#).
55. A. L. Read *Journal of Physics G: Nuclear and Particle Physics* **28** no. 10, (2002) 2693.
56. **ATLAS, CMS and LHCb collaborations** *LHCb-CONF-2012-017* (May, 2012) .
57. H. Nilles *Physics Reports* **110** no. 12, (1984) 1 – 162.
58. H. Haber and G. Kane *Physics Reports* **117** no. 24, (1985) 75 – 263.
59. M. Drees and M. M. Nojiri *Phys.Rev.* **D47** (1993) 376–408, [arXiv:hep-ph/9207234 \[hep-ph\]](#).
60. J. R. Ellis, K. A. Olive, and P. Sandick *Phys.Rev.* **D78** (2008) 075012, [arXiv:0805.2343 \[hep-ph\]](#).
61. F. Mahmoudi [arXiv:1205.3099 \[hep-ph\]](#). Proceedings of Moriond QCD 2012.
62. O. Buchmueller, R. Cavanaugh, A. De Roeck, M. Dolan, J. Ellis, *et al.* *Eur.Phys.J.* **C72** (2012) 1878, [arXiv:1110.3568 \[hep-ph\]](#).
63. **ATLAS collaboration**, F. Gianotti *CERN Seminar, July 4th, 2012* .
64. **CMS collaboration**, J. Incandela *CERN Seminar, July 4th, 2012* .
65. A. Arbey, M. Battaglia, A. Djouadi, and F. Mahmoudi [arXiv:1207.1348 \[hep-ph\]](#).
66. G. D’Ambrosio, G. Giudice, G. Isidori, and A. Strumia *Nucl.Phys.* **B645** (2002) 155–187, [arXiv:hep-ph/0207036 \[hep-ph\]](#).
67. W. Altmannshofer, A. J. Buras, S. Gori, P. Paradisi, and D. M. Straub *Nucl.Phys.* **B830** (2010) 17–94, [arXiv:0909.1333 \[hep-ph\]](#).

68. D. M. Straub [arXiv:1205.6094](#) [hep-ph].